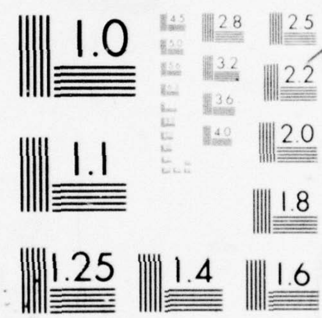
AD-A031 973 WISCONSIN UNIV MADISON MATHEMATICS RESEARCH CENTER F/G 12/1
A PRODUCT VERSION OF A VARIABLE METRIC METHOD AND ITS CONVERGEN--ETC(U)
SEP 76 K RITTER DAAG29-75-C-0024
UNCLASSIFIED MRC-TSR-1677 NL

1 OF 1
AD
A031973

END
DATE
FILMED
1-77



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

(11)

FG

MRC Technical Summary Report #1677

A PRODUCT VERSION OF A VARIABLE METRIC METHOD AND ITS CONVERGENCE PROPERTIES

Klaus Ritter

See 1473

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

September 1976

Received August 20, 1976

Approved for public release
Distribution unlimited

Sponsored by

U. S. Army Research Office
P.O. Box 12211
Research Triangle Park
North Carolina 27709

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

A PRODUCT VERSION OF A VARIABLE METRIC METHOD AND ITS CONVERGENCE PROPERTIES

Klaus Ritter

Technical Summary Report #1677
September 1976

ABSTRACT

An algorithm is described which uses a product representation of the matrix H_j approximating the inverse Hessian matrix of the function to be minimized. It is shown that the algorithm generates the same sequence of points as the Broyden-Fletcher-Goldfarb-Shanno- method. Using a simple relation between the traces of the matrices H_j and H_{j+1} corresponding to two consecutive points x_j and x_{j+1} the superlinear convergence of the algorithm is established.

AMS(MOS) Classification No. 90C30

Key Words: Nonlinear programming, unconstrained minimization, variable metric methods, superlinear convergence.

Work Unit No. 5 - Mathematical Programming and Operations Research

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024.

A PRODUCT VERSION OF A VARIABLE METRIC METHOD AND ITS CONVERGENCE PROPERTIES

Klaus Ritter

1. Introduction.

Various kinds of variable metric methods have been used effectively in the unconstrained minimization of a function of several variables. Their main feature is the use of rank one or two corrections to the matrix H which represents an approximation to the inverse Hessian matrix of the function to be minimized. Recently Brodlie, Gourlay and Greenstadt [1] and Davidon [3] used a representation of H as a product CC'. They showed that a rank two correction to H reduces to a rank one correction to C.

In this paper an algorithm is described which uses a simple rank one update formula for the matrix C. It is shown that the algorithm gives a product representation of the matrix H used in the Broyden-Fletcher-Goldfarb-Shanno-method [2], [4], [5], [6]. Under appropriate assumptions, the sequence generated by the algorithm converges superlinearly.

2. Formulation of the problem and notation.

Let $x \in E^n$ and let $F(x)$ be a real-valued function. We assume that $F(x)$ is twice continuously differentiable and denote the gradient and the Hessian matrix of $F(x)$ at a point x_i by $g_i = \nabla F(x_i)$ and $G_i = G(x_i)$, respectively. A prime is used for the transpose of a vector or a matrix.

We consider the problem of determining a z such that

$$(2.1) \qquad F(z) < F(x) \quad \text{for all} \quad x \neq z .$$

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024.

It is assumed that there are positive numbers μ, η and L such that

(2.2) $$\mu \|x\|^2 \leq x'G(y)x \leq \eta \|x\|^2 \qquad \text{for all } x, y \in E^n$$

and

$$\|G(x) - G(y)\| \leq L\|x-y\| \qquad \text{for all } x, y \in E^n .$$

Assumption (2.2) implies that $F(x)$ is uniformly convex and that there is a unique z with property (2.1). It is determined by the condition $\nabla F(z) = 0$. For later use we state two more well-known results which are immediate consequences of (2.2):

(2.3) $$\|G(y)\| \leq \eta , \quad \|(G(y)^{-1}\| \leq \mu^{-1} \qquad \text{for all } y \in E^n$$

(2.4) There are numbers $\gamma_1 > 0$ and $\gamma_2 > 0$ such that for all $x \in E^n$

$$\gamma_1 \|\nabla F(x)\| \leq \|x-z\| \leq \gamma_2 \|\nabla F(x)\| .$$

3. The Algorithm.

In a variable metric method at a given point x_j, the search direction s_j is determined by multiplying the gradient $g_j = \nabla F(x_j)$ by an appropriate matrix H_j, i.e.,

$$s_j = H_j g_j ,$$

where H_j is an approximation to the inverse Hessian matrix of $F(x)$ at x_j. With a suitable step size σ_j a new point

$$x_{j+1} = x_j - \sigma_j s_j$$

is computed. If $g_{j+1} = \nabla F(x_{j+1}) \neq 0$, the matrix H_{j+1} is determined from H_j in such a way that

(3.1) $$H_{j+1}(g_j - g_{j+1}) = \sigma_j s_j .$$

The various variable metric methods differ in the updating procedure which is used to compute H_{j+1} from H_j subject to (3.1). In many methods

H_j is a symmetric positive definite matrix and H_{j+1} is obtained by adding two symmetric matrices of rank one to H_j.

Any symmetric positive definite matrix H_j can be written in the form

$$H_j = C_j C_j' , \tag{3.2}$$

where C_j is a nonsingular matrix. Instead of adding a rank two correction matrix to H_j we can add a rank one matrix uv' to C_j and represent H_{j+1} in the form

$$H_{j+1} = (I + uv')CC'(I + vu') \tag{3.3}$$

where the vectors u and v are determined in such a way that (3.1) is satisfied and H_{j+1} is nonsingular.

Brodlie, Gourlay and Greenstadt [1] and more recently Davidon [3] have investigated representations of the form (3.2) and (3.3) of H_j and H_{j+1}, respectively, for some variable metric methods. In the following we shall describe an algorithm which uses a very simple update procedure of the form (3.3). It will be shown that this algorithm produces the same matrix H_{j+1} as the Broyden-Fletcher-Goldfarb-Shanno-method; see [2], [4], [5], [6].

We describe a general cycle of the algorithm. At the beginning of the jth cycle the following data is available: $C_j = (c_{1j}, \ldots, c_{nj})$, x_j, g_j and $\alpha_{ij} = c_{ij}' g_j$, $i = 1, \ldots, n$. For the initial cycle any nonsingular $(n \times n)$-matrix C_0 can be used.

<u>Step 1:</u> Computation of the direction of descent s_j.

Set

$$s_j = \sum_{i=1}^{n} \alpha_{ij} c_{ij}$$

Step 2: Computation of the step size σ_j.

Compute σ_j such that

$$F(x_j - \sigma_j s_j) = \min\{F(x_j - \sigma s_j) \mid \sigma \geq 0\}$$

Set

$$x_{j+1} = x_j - \sigma_j s_j$$

and compute g_{j+1}. If $g_{j+1} = 0$ stop, otherwise go to Step 3.

Step 3: Computation of C_{j+1}.

Compute

$$\beta_{ij} = c'_{ij} g_{j+1}, \qquad i = 1,\ldots,n,$$

$$\gamma_j = s'_j g_j, \qquad \delta_j = s'_j g_{j+1},$$

$$\omega_{ij} = \frac{1}{\gamma_j - \delta_j}[(\sqrt{(1 - \delta_j \gamma_j^{-1})\sigma_j} - 1)\alpha_{ij} + \beta_{ij}], \qquad i = 1,\ldots,n.$$

Set

$$c_{i,j+1} = c_{ij} + \omega_{ij} s_j \qquad i = 1,\ldots,n.$$

$$\alpha_{i,j+1} = \beta_{ij} + \omega_{ij} \delta_j$$

Replace j with j+1 and go to Step 1.

Remark:

In Step 1, the search direction s_j is determined by

$$s_j = C_j C'_j g_j.$$

In Step 2, we assume that σ_j is the optimal step size. This assumption is made in order to avoid some technical difficulties in the convergence proofs. The algorithm also works with an approximation to the optimal step size which after a finite number of steps can be set equal to 1. This will be shown in a much more general context in a subsequent paper. In Step 3, C_{j+1} is obtained by simply adding a multiple of s_j to each row of C_j. Obviously

$\delta_j = 0$, if σ_j is the optimal step size.

In the following theorem it will be shown that the search directions s_j generated by the above algorithm are identical with the search directions used in the Broyden-Fletcher-Goldfarb-Shanno-method.

Theorem 1:

Let $c_{1j}, \ldots, c_{nj}$, σ_j, s_j, g_j and g_{j+1} be defined by the algorithm. Set

$$C_j = (c_{ij}, \ldots, c_{nj}), \qquad H_j = C_j C_j',$$

$$p_j = \frac{s_j}{\|s_j\|}, \qquad d_j = \frac{g_j - g_{j+1}}{\|\sigma_j s_j\|}$$

and

$$H_{j+1} = H_j + \frac{d_j' p_j + d_j' H_j d_j}{(d_j' p_j)^2} p_j p_j' - \frac{p_j d_j' H_j + H_j d_j p_j'}{d_j' p_j}.$$

Then

$$s_{j+1} = C_{j+1} C_{j+1}' g_{j+1}$$

and

$$H_{j+1} = C_{j+1} C_{j+1}'.$$

Proof:

The first assertion follows immediately from Steps 1 and 3 of the algorithm since

$$c_{i,j+1}' g_{j+1} = (c_{ij} + \omega_{ij} s_j)' g_{j+1} = \beta_{ij} + \omega_{ij} \delta_j = \alpha_{i,j+1}.$$

In order to prove the second statement we observe that

$$c_{i,j+1}' - c_{ij}' = ((\sqrt{(1 - \delta_j \gamma_j^{-1})\sigma_j} - 1) \frac{c_{ij}' g_j}{s_j'(g_j - g_{j+1})} + \frac{c_{ij}' g_{j+1}}{s_j'(g_j - g_{j+1})}) s_j'$$

$$= (\sqrt{(1 - \delta_j \gamma_j^{-1})\sigma_j} \frac{c_{ij}' g_j}{p_j'(g_j - g_{j+1})} - \frac{c_{ij}'(g_j - g_{j+1})}{p_j'(g_j - g_{j+1})}) p_j.$$

Since

$$\sqrt{(1-\delta_j\gamma_j^{-1})\sigma_j}\ \frac{1}{p_j'(g_j-g_{j+1})} = \sqrt{\frac{\sigma_j p_j'(g_j-g_{j+1})}{p_j' g_j}}\ \frac{1}{p_j'(g_j-g_{j+1})}$$

$$= \sqrt{\frac{\sigma_j}{p_j' g_j\ p_j'(g_j-g_{j+1})}}$$

it follows that

$$C_{j+1}' - C_j' = \left(\frac{1}{\sqrt{p_j' g_j\ d_j' p_j}}\ \frac{C_j' g_j}{\sqrt{\|s_j\|}} - \frac{1}{d_j' p_j}\ C_j' d_j\right)p_j' \ .$$

Therefore

$$\begin{aligned}
C_{j+1}C_{j+1}' &= C_jC_j' + (C_{j+1}-C_j)C_j' + C_j(C_{j+1}'-C_j') + (C_{j+1}-C_j)(C_{j+1}'-C_j)\\
&= H_j + p_j\left(\frac{1}{\sqrt{p_j' g_j\ d_j' p_j}}\ \frac{s_j'}{\sqrt{\|s_j\|}} - \frac{d_j' H_j}{d_j' p_j}\right)\\
&\quad + \left(\frac{1}{\sqrt{p_j' g_j\ d_j' p_j}}\ \frac{s_j}{\sqrt{\|s_j\|}} - \frac{H_j d_j}{d_j' p_j}\right)p'\\
&\quad + p_j\left(\frac{g_j' H_j g_j}{d_j' p_j g_j' p_j \|s_j\|} - \frac{2\, d_j' H_j g_j}{d_j' p_j \sqrt{p_j' g_j d_j' p_j}\,\|s_j\|} + \frac{d_j' H_j d_j}{(d_j' p_j)^2}\right)p_j'\\
&= H_j + \frac{d_j' p_j + d_j' H_j d_j}{(d_j' p_j)^2}\, p_j p_j' - \frac{p_j\, d_j' H_j + H_j d_i p_j'}{d_j' p_j} \ .
\end{aligned}$$

4. Convergence.

If the algorithm terminates with some x_j, then $\nabla F(x_j) = 0$ and the assumptions on $F(x)$ imply that x_j is the unique global minimizer of $F(x)$. For the remainder of the paper we shall assume that the algorithm generates an infinite sequence $\{x_j\}$. In order to prove that it converges to the global minimizer of $F(x)$ we shall use the equivalence with the BFGS-Method established in Theorem 1.

First we shall write H_j as a sum of n matrices of rank one. For this purpose let

$$\rho_j = \|H_j g_j\|^{-1} \quad , \quad \lambda_j = \|H_j g_{j+1}\|^{-1}$$

and

$$H_j \rho_j g_j = p_j \quad , \quad H_j \lambda_j g_{j+1} = q_j \; .$$

Then $\|p_j\| = \|q_j\| = 1$ and $\rho_j g_j$ and $\lambda_j g_{j+1}$ are conjugate with respect to H_j since

$$\rho_j g_j' H_j g_{j+1} = p_j' g_{j+1} = 0 \; .$$

Assuming that H_j is positive definite we can find vectors $d_{3j}, \ldots, d_{nj}$ such that

$$H_j d_{ij} = p_{ij} \; , \quad \|p_{ij}\| = 1 \; , \qquad i = 3, \ldots, n$$

and

$$\rho_j g_j, \; \lambda_j g_{i+1}, \; d_{3j}, \ldots, d_{nj}$$

are conjugate with respect to H_j. Then

$$H_j = \frac{p_j p_j'}{\rho_j g_j' p_j} + \frac{q_j q_j'}{\lambda_j g_{j+1}' q_j} + \sum_{i=3}^{n} \frac{p_{ij} p_{ij}'}{d_{ij}' p_{ij}} \tag{4.1}$$

and

$$H_j^{-1} = \frac{\rho_j g_j g_j'}{g_j' p_j} + \frac{\lambda_j g_{j+1} g_{j+1}'}{g_{j+1}' q_j} + \sum_{i=3}^{n} \frac{d_{ij} d_{ij}'}{d_{ij}' p_{ij}} \; . \tag{4.2}$$

Using (4.1) we can easily derive a corresponding expression for H_{j+1}. First we observe that

$$H_j d_j = \frac{H_j g_j - H_j g_{j+1}}{\|\sigma_j s_j\|} = \frac{1}{\sigma_j} p_j - q_j \frac{\|H_j g_{j+1}\|}{\|\sigma_j s_j\|}$$

implies that

$$H_{j+1} x = H_j x \quad \text{for } x \in T_j = \{x \,|\, p_j' x = q_j' x = 0\} \; .$$

$d_{1j}, \ldots, d_{nj} \in T_j$ we have to change only the first two terms in (4.1) in order to obtain a representation of H_{j+1}, i.e., we have to determine two vectors in the span of g_j and g_{j+1} which are conjugate with respect to H_{j+1}.

We have

(4.3) $\quad d_j \in \text{span}\,\{g_j, g_{j+1}\}, \quad H_{j+1} d_j = p_j, \quad \text{and} \quad d_j' H_{j+1} d_j = d_j' p_j > 0 .$

Furthermore,

(4.4) $$g_{j+1}' H_{j+1} d_j = g_{j+1}' p_j = 0 .$$

Thus g_{j+1} and d_j are conjugate with respect to H_{j+1}. By the definition of H_{j+1}

(4.5) $$H_{j+1} g_{j+1} = H_j g_{j+1} - p_j \frac{d_j' H_j g_{j+1}}{d_j' p_j}$$

$$= (g_j - \alpha_j p_j) \| H_j g_{j+1} \| , \qquad \alpha_j = \frac{d_j' q_j}{d_j' p_j}$$

and

(4.6) $$g_{j+1}' H_{j+1} g_{j+1} = g_{j+1}' q_j \| H_j g_{j+1} \| > 0.$$

With

$$\rho_{j+1} = \| H_{j+1} g_{j+1} \|^{-1}, \quad H_{j+1} \rho_{j+1} g_{j+1} = p_{j+1}$$

it follows from (4.1)-(4.6), that H_{j+1} and H_{j+1}^{-1} are positive definite and can be written in the form

(4.7) $$H_{j+1} = \frac{p_{j+1} p_{j+1}'}{\rho_{j+1} g_{j+1}' p_{j+1}} + \frac{p_j p_j'}{d_j' p_j} + \sum_{i=3}^{n} \frac{p_{ij} p_{ij}'}{d_{ij}' p_{ij}}$$

and

(4.8) $$H_{j+1}^{-1} = \frac{\rho_{j+1} g_{j+1} g_{j+1}'}{g_{j+1}' p_{j+1}} + \frac{d_j d_j'}{d_j' p_j} + \sum_{i=1}^{n} \frac{d_{ij} d_{ij}'}{d_{ij}' p_j} .$$

By the definition of H_{j+2}

$$H_{j+2} x = H_{j+1} x \quad \text{for} \quad x \in T_{j+1} = \{x \mid H_{j+1} g_{j+1} = H_{j+1} g_{j+2} = 0\}.$$

Setting

$$q_{j+1} = \frac{H_{j+1} g_{j+2}}{\| H_{j+1} g_{j+2} \|} , \quad \lambda_{j+1} = \frac{1}{\| H_{j+1} g_{j+2} \|}$$

and observing that

$$\rho_{j+1} g_{j+1}' H_{j+1} g_{j+2} = p_{j+1}' g_{j+2} = 0$$

we can write H_{j+1} and H_{j+1}^{-1} as follows

$$(4.9)\quad H_{j+1} = \frac{p_{j+1}p'_{j+1}}{\rho_{j+1}g'_{j+1}p_{j+1}} + \frac{q_{j+1}q'_{j+1}}{\lambda_{j+1}g'_{j+2}q_{j+1}} + \sum_{i=3}^{n} \frac{p_{i,j+1}p'_{i,j+1}}{d'_{i,j+1}p_{i,j+1}}$$

and

$$(4.10)\quad H_{j+1}^{-1} = \frac{\rho_{j+1}g_{j+1}g_{j+1}}{g'_{j+1}p_{j+1}} + \frac{\lambda_{j+1}g_{j+2}g'_{j+2}}{g'_{j+2}q_{j+1}} + \sum_{i=3}^{n} \frac{d_{i,j+1}d'_{i,j+1}}{d'_{i,j+1}p_{i,j+1}}$$

where $d_{3,j+1},\ldots,d_{n,j+1}$ are vectors in T_{j+1} with

$$d'_{i,j+1}\; H_{j+1}\; d_{k,j+1} = 0\;,\quad i \neq k\;,\quad i,k = 3,\ldots,n\;,$$

$$\|H_{j+1}d_{i,j+1}\| = 1\;,\quad H_{j+1}d_{i,j+1} = p_{i,j+1},\quad i = 3,\ldots,n.$$

This representation is completely analoguous to (4.1) and (4.2) and can be used to derive a representation of H_{j+2} as a sum of n matrices of rank one. In the following convergence proof we shall use a simple argument involving the trace of the matrices H_j and H_{j+1}. By definition, the trace of a square matrix M, denoted by tr(M), is the sum of the diagonal elements of M. From (4.1) and (4.2) we have

$$(4.11)\quad tr(H_j) = tr\left(\frac{p_jp'_j}{\rho_jg'_jp_j}\right) + tr\left(\frac{q_jq'_j}{\lambda_jg'_{j+1}q_j}\right) + \sum_{i=3}^{n} tr\left(\frac{p_{ij}p_{ij}}{d'_{ij}p_{ij}}\right)$$

$$= \frac{1}{\rho_j\, g'_j\, p_j} + \frac{1}{\lambda_j\, g'_{j+1}q_j} + \sum_{i=3}^{n} \frac{1}{d'_{ij}\, p_{ij}}$$

and

$$(4.12)\quad tr(H_j^{-1}) = \frac{\rho_i\|g_j\|^2}{g'_j\; p_j} + \frac{\lambda_j\|g_{j+1}\|^2}{g'_{j+1}\,q_j} + \sum_{i=3}^{n} \frac{\|d_{ij}\|^2}{d'_{ij}\;p_{ij}}\;.$$

Setting $\tau_j = (1+\lambda_j^2\|g_{j+1}\|^2)/\lambda_j\, g'_{j+1}\, q_j,\ \omega_j = (1+\|d_j\|^2)/d_jp_j$,

$$(4.13)\qquad \varphi_j = \tau_j + \sum_{i=3}^{n} \frac{1+\|d_{ij}\|^2}{d'_{ij}\;p_{ij}}\quad ,$$

and

$$(4.15) \qquad \varphi_{j+1} = \frac{1 + \lambda_{j+1}^2 \|g_{j+2}\|^2}{\lambda_{j+1} g'_{j+2} q_{j+1}} + \sum_{i=3}^{n} \frac{1 + \|d_{i,j+1}\|^2}{d'_{i,j+1} p_{i,j+1}}$$

we conclude from (4.7) - (4.12) that

$$(4.16) \qquad \varphi_{j+1} = \mathrm{tr}(H_{j+1}) - \frac{1}{\rho_{j+1} g'_{j+1} p_{j+1}} + \mathrm{tr}(H^{-1}_{j+1}) - \frac{\rho_{j+1} \|g_{j+1}\|^2}{g'_{j+1} p_{j+1}}$$

$$= \omega_j + \sum_{i=3}^{n} \frac{1 + \|d_{ij}\|^2}{d'_{ij} p_{ij}}$$

$$= \varphi_j - \tau_j + \omega_j .$$

The equality (4.16) will be used to prove that the sequence of gradients produced by the algorithm converges to zero. First we observe that $\varphi_{j+1} > 0$ since H_{j+1} is positive definite and, therefore, all terms on the right hand side of (4.15) are positive. Secondly we shall show (Lemma 1) that, because of the assumptions on $F(x)$, the sequence $\{\omega_j\}$ is bounded. This implies that there is a constant γ and an infinite set $J \subset \{0,1,2,\ldots\}$ such that

$$(4.17) \qquad \tau_j \leq \gamma \quad \text{for } j \in J .$$

Using the definition of τ_j and p_{j+1} it can easily be shown (Lemma 2) that there is $\varepsilon > 0$ such that

$$g'_{j+1} p_{j+1} \geq \varepsilon \|g_{j+1}\| \quad \text{for} \quad j \in J .$$

By a routine argument it follows from this that

$$\|g_{j+1}\| \to 0 \text{ as } j \to \infty , \quad j \in J ,$$

which by the uniform convexity of $F(x)$ implies

$$\|g_{j+1}\| \to 0 \quad \text{as } j \to \infty$$

<u>Lemma 1.</u>

$$\omega_j = \frac{1 + \|d_j\|^2}{d'_j p_j} \leq \frac{1 + \eta^2}{\mu} \qquad \text{for all } j .$$

Proof:

By Taylor's theorem there are

$$\xi_j, \eta_j \;\epsilon\; \{x \mid x = x_j - t(\sigma_j s_j),\ 0 \leq t \leq 1\}$$

such that

$$p'_j\, g_{j+1} = p'_j g_j - \sigma_j\, p'_j G(\xi_j) s_j$$

and

$$d'_j g'_{j+1} = d'_j\, g_j - \sigma_j\, d'_j\, G(\eta_j) s_j\ .$$

Therefore,

$$d'_j\, p_j = \frac{(g_j - g_{j+1})'\, p_j}{\|\sigma_j s_j\|} = p'_j\, G(\xi_j) p_j \geq \mu \tag{4.18}$$

and

$$\|d_j\|^2 = \frac{d'_j(g_j - g_{j+1})}{\|\sigma_j s_j\|} = d'_j G(\eta_i) p_j \leq \|d_j\|\,\|G(\eta_j)\| \leq \eta \|d_j\|. \tag{4.19}$$

Lemma 2.

For every $\gamma > 0$ there is $\varepsilon > 0$ such that, for all j,

$$\tau_j = \frac{1 + \lambda_j^2 \|g_{j+1}\|^2}{\lambda_j\, g'_{j+1}\, q_j} \leq \gamma \ \text{ implies } \ g'_{j+1} p_{j+1} \geq \varepsilon \|g_{j+1}\|\ .$$

Proof:

If $\tau_j \leq \gamma$, then

$$\gamma \geq \frac{\|g_{j+1}\|}{g'_{j+1} q_j}\left(\frac{1}{\lambda_j \|g_{j+1}\|} + \lambda_j \|g_{j+1}\|\right) \geq 2\ \frac{\|g_{j+1}\|}{g'_{j+1} q_j}$$

and

$$\frac{g'_{j+1} q_j}{\|g_{j+1}\|} \geq \frac{2}{\gamma}\ .$$

By (4.5),

$$p_{j+1} = (q_j - \frac{d'_j q_j}{d'_j p_j}\, p_j) \|q_j - \frac{d'_j q_j}{d_j p_j}\, p_j\|^{-1}\ .$$

Since, by (4.18) and (4.19),

$$\|q_j - \frac{d_j'q_j}{d_j'p_j} p\| \leq 1 + \frac{|d_j q_j|}{d_j'p_j} \leq 1 + \frac{\|d_j\|}{d_j'p_j} \leq 1 + \frac{\eta}{\mu} ,$$

it follows that

$$\frac{p_{j+1}' \, g_{j+1}}{\|g_{j+1}\|} \geq \frac{\mu}{1+\eta} \, \frac{g_{j+1}' \, q_j}{\|g_{j+1}\|} \geq \frac{2\mu}{\gamma(1+\eta)}.$$

Theorem 2.

The sequence $\{x_j\}$ generated by the algorithm converges to a z such that

$$F(z) < F(x) \quad \text{for all} \quad x \neq z$$

and

$$\nabla F(z) = 0 .$$

Proof:

By (4.17) and Lemma 1 and 2 there is $\varepsilon > 0$ and an infinite set $J \subset \{0,1,2,\dots\}$ such that

$$g_j'p_j \geq \varepsilon \|g_j\| \quad \text{for } j \in J.$$

Furthermore, it follows from Taylor's theorem that, for some $\xi_j \in \{x \mid x = x_j - t(\sigma s_j), 0 \leq t \leq 1\}$,

$$\begin{aligned} F(x_j - \sigma s_j) - F(x_j) &= -\sigma g_j's_j + \tfrac{1}{2}\sigma^2 s_j'G(\xi_j)s_j \\ &\leq -\sigma g_j's_j + \tfrac{1}{2}\sigma^2\eta\|s\|^2 . \end{aligned}$$

Thus choosing

$$\tilde{\sigma} = \frac{g_j' s_j}{\eta\|s_j\|^2}$$

we obtain, for $j \in J$,

$$F(x_{j+1}) - F(x_j) \le F(x_j - \tilde{\sigma} s_j) - F(x_j)$$

$$\le -\frac{1}{2\eta} \frac{(g_j' s_j)^2}{\|s_j\|^2}$$

$$\le -\frac{\varepsilon^2}{2\eta} \|g_j\| .$$

Since $F(x)$ is bounded from below and $F(x_{j+1}) < F(x_j)$ for all j, it follows that

$$\|g_j\| \to 0 \text{ as } j \to \infty , \ j \in J .$$

Let z be a cluster point of the sequence $\{x_j, j \in J\}$. Then $\nabla F(z) = 0$. By the uniform convexity of $F(x)$, z is the unique global minimizer of $F(x)$. Since $F(x_{j+1}) < F(x_j)$ for all j, the sequence $\{x_j\}$ has no cluster point except z. Thus $x_j \to z$ as $j \to \infty$.

5. Superlinear convergence.

In order to prove that

$$\frac{\|g_{j+1}\|}{\|g_j\|} \to 0 \quad \text{as} \quad j \to \infty$$

we shall again use the trace of the matrices H_j and H_{j+1}. The argument will be based on a modification of the formula (4.16).

By Theorem 2, $x_j \to z$ as $j \to \infty$. Let $G = G(z)$ and denote by $G^{\frac{1}{2}}$ the square root of G. By definition $G^{\frac{1}{2}}$ is a symmetric positive definite matrix with the property

$$G = G^{\frac{1}{2}} G^{\frac{1}{2}} .$$

We set $G^{-\frac{1}{2}} = (G^{\frac{1}{2}})^{-1}$. Replacing the matrices H_j and H_j^{-1} by

$$G^{\frac{1}{2}} H_j G^{\frac{1}{2}} \quad \text{and} \quad (G^{\frac{1}{2}} H_j G^{\frac{1}{2}})^{-1} = G^{-\frac{1}{2}} H_j^{-1} G^{-\frac{1}{2}} ,$$

respectively, in the formulae (4.1), (4.2) and (4.7) - (4.16) we obtain

$$\omega_j = \frac{p_j' G p_i + d_j' G^{-1} d_j}{d_j' p_j}, \qquad \tau_j = \frac{q_j' G q_j + \lambda_j^2 g_{i+1} G^{-1} g_{j+1}}{\lambda_j g_{j+1}' q_j},$$

$$\varphi_j = \tau_j + \sum_{i=3}^{n} \frac{p_{ij}' G p_{ij} + d_{ij}' G^{-1} d_{ij}}{d_{ij}' p_{ij}}$$

and

(5.1) $$\varphi_{j+1} = \zeta_j - \tau_j + \omega_j .$$

In order to prove that $\{\|g_j\|\}$ converges superlinearly to zero we observe that by Taylor's theorem

$$\begin{aligned}\|d_j - Gp_j\| &= \|(\int_0^1 G(x_j - t\sigma_j s_j)dt - G)p_j\| \\ &\le \sup_{0 \le t \le 1} \|G(x_j - t\sigma_j s_j) - G_j\| + \|G_j - G\| \\ &\le L\|x_j - x_{j+1}\| + L\|x_j - z\| \\ &= 0(\max\{\|g_j\|, \|g_{j+1}\|\}) .\end{aligned}$$

Thus Lemma 3 below implies that $\omega_j \to 2$ as $j \to \infty$. Since $\varphi_j > 0$ we shall conclude from this that $\tau_j \to 2$ as $j \to \infty$. Setting

$$\psi_j = \mathrm{tr}(\frac{G^{\frac{1}{2}} p_j p_j' G^{\frac{1}{2}}}{\rho_j g_j' p_j}) + \mathrm{tr}(\frac{\rho_j G^{-\frac{1}{2}} g_j g_j' G^{-\frac{1}{2}}}{g_j' p_j})$$

we obtain from Lemma 4 that $\psi_j \to 2$ as $j \to \infty$ which by Lemma 5 in turn implies that $\|g_{j+1}\| \|g_j\|^{-1} \to 0$ as $j \to \infty$.

<u>Lemma 3.</u>

Let $x, y \in E^n$ be such that $\|x\| = 1$ and $y'x > 0$. Set

$$\tau = \mathrm{tr}(\frac{G^{\frac{1}{2}} x x' G^{\frac{1}{2}}}{y'x}) + \mathrm{tr}(\frac{G^{-\frac{1}{2}} y y' G^{-\frac{1}{2}}}{y'x}).$$

Then

$$\tau \ge 2$$

and there is a constant $\gamma > 0$ such that $\tau - 2 \le \gamma$ implies

$$y'x \geq \tfrac{1}{2}\mu \,, \qquad \|Gx-y\|^2 \leq \eta(\eta+\mu)(\tau-2)$$

and $\|Gx - y\| \leq \gamma$ implies

$$\tau - 2 \leq \frac{2}{\mu^2}\|Gx - y\|^2 \,.$$

Proof:

Set $v = y - Gx$. Then

$$\tau = \frac{x'Gx + y'G^{-1}y}{y'x} = \frac{2x'Gx + 2v'x + v'G^{-1}v}{x'Gx + v'x}$$

$$= 2 + \frac{vG^{-1}v}{x'Gx + v'x}$$

Thus $\tau \geq 2$ and

$$\frac{1}{\eta}\|v\|^2 \leq v'G^{-1}v = (\tau-2)(x'Gx + v'x)$$

$$\leq (\tau-2)(\|G\| + \|v\|) \leq (\tau-2)(\eta + \|v\|)$$

Therefore, we have

$$\|v\| \leq (\tau - 2)\eta\left[\frac{\eta}{\|v\|} + 1\right]$$

which implies that for $\tau-2$ sufficiently small $\|v\| \leq 0.5\mu$. Thus, for γ sufficiently small,

$$\|v\|^2 \leq \eta(\eta+\mu)(\tau-2)$$

and

$$y'x = x'Gx + v'x \geq \mu - \|v\| \geq 0.5\mu \,.$$

Finally, for $\|Gx-y\| \leq 0.5\mu$,

$$\tau-2 \leq \frac{\|v\|^2\|G^{-1}\|}{\mu - \|v\|} \leq \frac{2\|v\|^2}{\mu^2} \,.$$

Lemma 4.

There is a constant $\gamma > 0$ such that, for all j , $\tau_j - 2 \leq \gamma$ implies

$$|\psi_{j+1} - \tau_j| = 0(\max\{\sqrt{\tau_j-2}\,,\ \|g_j\|\}) \,.$$

Proof:

By (4.5)

$$p_{j+1} = (q_j - \alpha_j p_j)\|q_j - \alpha_j p_j\|^{-1} .$$

Therefore, $\rho_{j+1} = \lambda_{j+1}\|q_j - \alpha_j p_j\|^{-1}$ implies

$$\rho_{j+1} g'_{j+1} p_{j+1} = \rho_{j+1} g'_{j+1} q_j \|q_j - \alpha_j p_j\|^{-1}$$

$$= \lambda_{j+1} g'_{j+1} q_j \|q_j - \alpha_j p_j\|^{-2} .$$

Furthermore,

$$p'_{j+1} G p_{j+1} = \|q_j - \alpha_j p_j\|^{-2} (q_j - \alpha_j p_j)' G (q_j - \alpha_j p_j)$$

and

$$\psi_{j+1} = \frac{p'_{j+1} G p_{j+1} + \rho^2_{j+1} g'_{j+1} G^{-1} g_{j+1}}{\rho_{j+1} g'_{j+1} p_{j+1}}$$

$$= \frac{(q_j - \alpha_j p_j)' G (q_j - \alpha_j p_j) + \rho^2_{j+1}\|q_j - \alpha_j p_j\|^2 g'_{j+1} G^{-1} g_{j+1}}{\lambda_{j+1} g'_{j+1} q_j}$$

$$= \tau_j + \frac{\alpha_j^2 p_j' G p_j - 2\alpha_j q_j' G p_j}{\lambda_{j+1} g'_{j+} q_j} .$$

By definition, $\alpha_j = d_j' q_j / d_j' p_j$ and

$$d_j' q_j = p_j' G q_j + (d_j' - p_j' G) q_j$$

$$= p_j'(\lambda_{j+1} g_{j+1}) - p_j'(\lambda_{j+1} g_{j+1} - G q_j) + (d_j' - p_j' G) q_j .$$

Therefore,

$$|d_j' q_j| \leq \|\lambda_{j+1} g_{j+1} - G q_j\| + \|d_j - G p_j\| .$$

Since, $\|d_j - G p_j\| = 0(\max\{\|g_j\|, \|g_{j+1}\|\}$, $d_j' p_j \geq \mu$ and Lemma 3 implies that for γ sufficiently small

$$\|\lambda_{j+1} g_{j+1} - G q_j\| = 0(\sqrt{\tau_j - 2})$$

and

$$\lambda_{j+1} g'_{j+1} q_j \geq 0.5\mu$$

we have

$$|\psi_{j+1} - \tau_j| \leq \frac{\alpha_j^2 \|G\| + 2\alpha_j \|G\|}{0.5\mu}$$

$$= 0(\max\{\sqrt{\tau_j - 2}, \|g_j\|, \|g_{j+1}\|\}).$$

Lemma 5.

There is a constant $\gamma > 0$ such that, for all j, $\psi_j - 2 \leq \gamma$ implies

i) $|\sigma_j - 1| = 0(\max\{\sqrt{\psi_j - 2}, \|g_j\|, \|g_{j+1}\|\})$

ii) $\frac{\|g_{j+1}\|}{\|g_j\|} = 0(\max\{\sqrt{\psi_j - 2}, \|g_j\| \|g_{j+1}\|\}).$

Proof:

By Taylor's theorem there is

$$\xi_j \in \{x | x = x_j - t(\sigma_j s_j), \ 0 \leq t \leq 1\}$$

such that

$$0 = s'_j g_{j+1} = s'_j g_j - \sigma_j s'_j G(\xi_j) s_j .$$

Thus

$$\sigma_j = \frac{g'_j s_j}{s'_j G(\xi_j) s_j} = 1 - \frac{g'_j s_j - s'_j G(\xi_j) s_j}{s'_j G(\xi_j) s_j} . \tag{5.2}$$

Since with $v_j = \rho_j g_j - G p_j$ we have

$$g'_j s_j = (\frac{p'_j}{\rho_j} G + \frac{v'_j}{\rho_j}) s_j = s'_j G s_j + v'_j s_j \|s_j\| ,$$

and it follows that

$$|\sigma_j - 1| = \frac{s_j' G s_j + v_j' s_j \|s_j\| - s_j' G(\xi_j) s_j}{s_j' G(\xi_j) s_j}$$

$$\leq \frac{s_j'(G - G(\xi_j)) s_j + \|v_j\| \|s_j\|^2}{\mu \|s_j\|^2}$$

$$\leq \frac{1}{\mu}(\|G - G(\xi_j)\| + \|v_j\|)$$

$$= 0(\max\{\sqrt{\psi_j - 2}, \|g_j\|, \|g_{j+1}\|),$$

where the last equality follows from Lemma 3 and (2.4) since

$$\|G - G(\xi_j)\| \leq \|G - G_j\| + \|G_j - G(\xi_j)\| \leq L\|x_j - z\| + L\|x_{j+1} - x_j\| .$$

Setting

$$E_j = \int_0^1 G(x_j - t\sigma_j s_j)dt - G$$

we have again by Taylor's theorem

$$g_{j+1} = g_j - \sigma_j G s_j - \sigma_j E_j s_j$$

$$= g_j - G s_j + (1 - \sigma_j) G s_j - \sigma_j E_j s_j$$

and

$$\frac{\|g_{j+1}\|}{\|g_j\|} \leq \frac{\|g_j - G s_j\|}{\|g_j\|} + |1 - \sigma_j| \frac{\|G s_j\|}{\|g_j\|} + \|E_j\| \frac{\|\sigma_j s_j\|}{\|g_j\|}$$

$$= \frac{\|\rho_j g_j - G p_j\|}{\rho_j \|g_j\|} + |1 - \sigma_j| \frac{\|G p_j\|}{\rho_j \|g_j\|} + \|E_j\| \frac{\|\sigma_j s_j\|}{\|g_j\|} .$$

Since $\|G p_j\| \geq \mu$ we obtain from Lemma 3, that for γ sufficiently small

$$\|\rho_j g_j - G p_j\| = 0(\sqrt{\psi_j - 2}) \quad \text{and} \quad \|\rho_j g_j\| \geq 0.5\mu .$$

Furthermore, by (5.2)

$$\frac{\|\sigma_j s_j\|}{\|g_j\|} \leq \frac{\|s_j\|}{\|g_j\|} \frac{\|g_j\| \|s_j\|}{\mu \|s_j\|^2} \leq \frac{1}{\mu}$$

and by (2.4),

$$\|E_j\| \leq \|\int_0^1 G(x_j - t\sigma_j s_j)dt - G_j\| + \|G_j - G\|$$

$$\leq \max_{0 \leq t \leq 1} \{\|G(x_j - t\sigma_j s_j) - G_j\| + \|G_j - G\|$$

$$\leq L\|x_{j+1} - x_j\| + L\|x_j - z\| = 0(\max\{\|g_j\|, g_{j+1}\|\}).$$

Therefore we have

$$\frac{\|g_{j+1}\|}{\|g_j\|} = 0(\max\{\sqrt{\psi_j - 2}, \|g_j\|, \|g_{j+1}\|\}) .$$

Theorem 3.

Let the sequences $\{\sigma_j\}$, $\{g_j\}$ and $\{x_j\}$ be generated by the algorithm. Then

i) $|\sigma_j - 1| \to 0$ as $j \to \infty$

ii) $\frac{\|g_{j+1}\|}{\|g_j\|} \to 0$ as $j \to \infty$

iii) $\frac{\|x_{j+1} - z\|}{\|x_j - z\|} \to 0$ as $j \to \infty$.

Proof:

By Theorem 2, $\|g_j\| \to 0$ as $j \to \infty$. If $\tau_j \to 2$ as $j \to \infty$, it follows from Lemma 4 that $\psi_j \to 2$, which by Lemma 5 and (2.4) implies the statements of the theorem. Thus it suffices to show that

$$\tau_j \to 2 \quad \text{as} \quad j \to \infty.$$

Since $\tau_j \geq 2$ this is equivalent with proving that for every $\varepsilon > 0$ there is $j(\varepsilon)$ such that

$$\tau_j \leq 2 + \varepsilon \quad \text{for} \quad j \geq j(\varepsilon) .$$

Since $\|d_j - Gp_j\| \to 0$ as $j \to \infty$, it follows from Lemma 3, that for j sufficiently large,

(5.3) $$\omega_j = 2 + 0(\|d_j - G p_j\|)$$
$$= 2 + 0(\max\{g_j\|, \|g_{j+1}\|\}).$$

Therefore, by (5.1)

$$\varphi_{j+1} = \varphi_j - \tau_j + 2 + 0(\max\{\|g_j\|, \|g_{j+1}\|\}).$$

Let

$$J = \{ j \mid \tau_j > 2 + \varepsilon \}.$$

Since $\|g_j\| \to 0$ as $j \to \infty$, there is j_1 such that

(5.4) $$\varphi_{j+1} \le \varphi_j - (2+\varepsilon) + 2 + 0(\max\{\|g_j\|, \|g_{j+1}\|\})$$
$$\le \varphi_j - \frac{\varepsilon}{2} \qquad \text{for } j \ge j_1 \quad \text{and } j \in J.$$

If $j \notin J$ and ε is sufficiently small it follows from Lemma 4 and 5 that there is j_2 such that

(5.5) $$\|g_{j+2}\| \le 0.5\|g_{j+1}\| \qquad \text{for} \quad j \ge j_2 \quad \text{and } j \notin J.$$

Thus, if $j - 1 \in J$, $j + i \notin J$, $i = 0,\ldots,k-1$, $j + k \in J$ it follows from (5.3) and (5.5) that for $j \ge j_2$

(5.6) $$\varphi_{j+k} - \varphi_j = (\varphi_{j+1} - \varphi_j) + (\varphi_{j+2} - \varphi_{j+1}) + \ldots + (\varphi_{j+k} - \varphi_{j+k-1})$$
$$= 0(\|g_j\|) + 0(\|g_{j+1}\|) + \ldots + 0(\|g_{j+k-1}\|)$$
$$= 0(\|g_j\|) + 0(2\|g_{j+1}\|).$$
$$\le \frac{\varepsilon}{4} \qquad \text{for } j \text{ sufficiently large}.$$

Since $\varphi_j > 0$, (5.4) and (5.6) imply that J has at most finitely many elements.

As a further application of (5.1) we obtain

Theorem 4.

The sequences

$$\{\|H_j\|\} \qquad \text{and} \qquad \{\|H_j^{-1}\|\}$$

are bounded.

Proof:

From (5.1) we have

$$\begin{aligned}\varphi_j &= \varphi_0 + \sum_{i=1}^{j-1} (\varphi_i - \varphi_{i-1}) \\ &= \varphi_0 + \sum_{i=1}^{j-1} (\omega_i - \tau_i) \\ &\leq \varphi_0 + \sum_{i=1}^{j-1} (\omega_i - 2)\end{aligned}$$

because $2 - \tau_i \leq 0$ for all i. Since $\| d_j - G\, p_j \| \to 0$ as $j \to \infty$ it follows from Lemma 3 that

$$\sum_{i=1}^{j-1} (\omega_i - 2) = 0\Big(\sum_{i=1}^{j-1} \| g_i \|^2\Big) = 0\Big(\sum_{i=1}^{\infty} \| g_i \|^2\Big) < \infty .$$

Thus the sequence $\{\varphi_j\}$ is bounded. Moreover, in the proof of the previous theorem it has been shown that $\tau_i \to 2$ as $j \to \infty$. Thus Lemma 4 implies that $\psi_i \to 2$ as $j \to \infty$. Therefore

$$0 < \mathrm{tr}(G^{\frac{1}{2}} H_j G^{\frac{1}{2}}) \;\psi\; \mathrm{tr}(G^{-\frac{1}{2}} H_j^{-1} G^{-\frac{1}{2}}) = \varphi_j + \psi_j$$

is bounded. Since the trace of a matrix is equal to the sum of its eigenvalues, this means that there is a uniform upper bound for the eigenvalues of the matrices $G^{\frac{1}{2}} H_j G^{\frac{1}{2}}$ and $G^{-\frac{1}{2}} H_j^{-1} G^{-\frac{1}{2}}$, $j = 0,1,2,\ldots$. Thus there is a uniform upper bound for the numbers $\| H_j \|$ and $\| H_j^{-1} \|$, $j = 0,1,2,\ldots$.

References

[1] K. W. Brodlie, A. R. Gourlay and J. Greenstadt, "Rank one and two corrections to positive definite matrices expressed in product form", Journal of the Institute of Mathematics and its Applications 11 (1973) 73-82.

[2] C. G. Broyden, "The convergence of a class of double-rank minimization algorithms", Journal of the Institute of Mathematics and its Applications 6 (1970) 76-90.

[3] W. C. Davidon, "Optimally conditioned optimization algorithms without line searches", Mathematical Programming 3 (1975) 1-30.

[4] R. Fletcher, "A new approach to variable metric algorithms", The Computer Journal 13 (1970) 23-26.

[5] D. Goldfarb, "A family of variable metric methods derived by variational means", Mathematics of Computation 24 (1970) 23-26.

[6] D. F. Shanno, "Conditioning of Quasi-Newton methods for function minimization", Mathematics of Computation 24 (1970) 647-656.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 1677

2. GOVT ACCESSION NO.:

3. RECIPIENT'S CATALOG NUMBER:

4. TITLE (and Subtitle): 6 A PRODUCT VERSION OF A VARIABLE METRIC METHOD AND ITS CONVERGENCE PROPERTIES.

5. TYPE OF REPORT & PERIOD COVERED: Summary Report - no specific reporting period

6. PERFORMING ORG. REPORT NUMBER:

7. AUTHOR(s): 10 Klaus Ritter

8. CONTRACT OR GRANT NUMBER(s): 15 DAAG29-75-C-0024

9. PERFORMING ORGANIZATION NAME AND ADDRESS:
Mathematics Research Center, University of Wisconsin
610 Walnut Street
Madison, Wisconsin 53706

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS:

11. CONTROLLING OFFICE NAME AND ADDRESS:
U. S. Army Research Office
P.O. Box 12211
Research Triangle Park, North Carolina 27709

12. REPORT DATE: 11 September 1976

13. NUMBER OF PAGES: 22

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): 9 Technical summary rept.,

15. SECURITY CLASS. (of this report): UNCLASSIFIED

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT (of this Report): 12 25p.

Approved for public release; distribution unlimited.

14 MRC-TSR-1677

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report):

18. SUPPLEMENTARY NOTES:

19. KEY WORDS (Continue on reverse side if necessary and identify by block number):
Nonlinear programming, unconstrained minimization, variable metric methods, superlinear convergence

20. ABSTRACT (Continue on reverse side if necessary and identify by block number):

A product version of a variable metric method is described. It is shown that it converges superlinearly.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED 221 200

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered) LB